

# UC San Diego

## UC San Diego Previously Published Works

### Title

Gut microbiota, short chain fatty acids, and obesity across the epidemiologic transition: the METS-Microbiome study protocol.

### Permalink

<https://escholarship.org/uc/item/189289g9>

### Journal

BMC public health, 18(1)

### ISSN

1471-2458

### Authors

Dugas, Lara R  
Lie, Louise  
Plange-Rhule, Jacob  
et al.

### Publication Date

2018-08-01

### DOI

10.1186/s12889-018-5879-6

Peer reviewed

STUDY PROTOCOL

Open Access



# Gut microbiota, short chain fatty acids, and obesity across the epidemiologic transition: the METS-Microbiome study protocol

Lara R. Dugas<sup>1\*</sup> , Louise Lie<sup>1</sup>, Jacob Plange-Rhule<sup>2</sup>, Kweku Bedu-Addo<sup>2</sup>, Pascal Bovet<sup>3,4</sup>, Estelle V. Lambert<sup>5</sup>, Terrence E. Forrester<sup>6</sup>, Amy Luke<sup>1</sup>, Jack A. Gilbert<sup>7</sup> and Brian T. Layden<sup>8,9</sup>

## Abstract

**Background:** While some of the variance observed in adiposity and weight change within populations can be accounted for by traditional risk factors, a new factor, the gut microbiota, has recently been associated with obesity. However, the causal mechanisms through which the gut microbiota and its metabolites, short chain fatty acids (SCFAs) influence obesity are unknown, as are the individual obesogenic effects of the individual SCFAs (butyrate, acetate and propionate). This study, METS-Microbiome, proposes to examine the influence of novel risk factors, the gut microbiota and SCFAs, on obesity, adiposity and weight change in an international established cohort spanning the epidemiologic transition.

**Methods:** The parent study; Modeling the Epidemiologic Transition Study (METS) is a well-established and ongoing prospective cohort study designed to assess the association between body composition, physical activity, and relative weight, weight gain and cardiometabolic disease risk in five diverse population-based samples in 2500 people of African descent. The cohort has been prospectively followed since 2009. Annual measures of obesity risk factors, including body composition, objectively measured physical activity and dietary intake, components which vary across the spectrum of social and economic development. In our new study; METS-Microbiome, in addition to continuing yearly measures of obesity risk, we will also measure gut microbiota and stool SCFAs in all contactable participants, and follow participants for a further 3 years, thus providing one of the largest gut microbiota population-based studies to date.

**Discussion:** This new study capitalizes upon an existing, extensively well described cohort of adults of African-origin, with significant variability as a result of the widespread geographic distributions, and therefore variation in the environmental covariate exposures. The METS-Microbiome study will substantially advance the understanding of the role gut microbiota and SCFAs play in the development of obesity and provide novel obesity therapeutic targets targeting SCFAs producing features of the gut microbiota.

**Trial registration:** Registered [NCT03378765](https://clinicaltrials.gov/ct2/show/study/NCT03378765) Date first posted: December 20, 2017.

**Keywords:** Epidemiologic transition, Gut microbiota, Short chain fatty acids, Obesity

\* Correspondence: [ldugas@luc.edu](mailto:ldugas@luc.edu)

<sup>1</sup>Public Health Sciences, Stritch School of Medicine, Loyola University  
Chicago, 2160 S. 1st Avenue, Maywood, IL 60153, USA

Full list of author information is available at the end of the article



© The Author(s). 2018 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

## Background

Obesity is a complex condition with a multi-faceted etiology. To date, while some of the variance observed in obesity within populations can be accounted for by traditional risk factors such as total energy expenditure, physical activity (PA) level, dietary intake, genetics, socioeconomic status or education level [1–6], a new factor, the gut microbiota have also been recently implicated in obesity [7–9]. This study, METS-Microbiome, proposes to examine the influence of the gut microbiota, and its metabolites of carbohydrate digestion, short chain fatty acids (SCFAs), on obesity, and weight change.

The gut microbes are responsible for breaking down non-digestible dietary nutrients, such as pectin, cellulose, and resistant starches. Fermentation of these nutrients in the distal gut results in the production of SCFAs, mainly butyrate, propionate and acetate. Each of these is absorbed by the human gut and contributes approximately 200 kcal/day to total body energy expenditure [10]. SCFAs are a key energy source for the intestinal epithelium and liver [11], and consequently affect many metabolically important processes including hepatic gluconeogenesis and lipogenesis [12, 13], gut barrier function [14, 15], gut motility [16] and immune responses [17, 18]. Importantly, digestion of resistant starches, with associated increases in fecal SCFA concentrations, has been shown to increase satiety, and is associated with improvements in blood glucose and cholesterol [19, 20].

There are approximately 39 trillion bacterial cells in the human microbiota in an average-sized man, whereas there are 30 trillion human cells in the body [21]. Currently bacteria can be classified in 29 phyla and 5569 taxa, along the List of Prokaryotic names with Standing in Nomenclature [22]. Three main phyla colonize the gut; Firmicutes, Bacteroidetes and less abundantly by the Actinobacteria. In humans, studies indicate an increase in the Firmicutes and a decrease in the Bacteroidetes phyla to be associated with obesity [23, 24], although not all studies have observed this [25, 26]. In one of the earliest human studies, Ley et al. [23] compared the gut microbiota of 12 obese individuals, following two different low calorie diets over the period of 1 year and found that at baseline, obesity was associated with fewer Bacteroidetes ( $p < 0.001$ ). However, with subsequent weight loss, there were increases in the Bacteroidetes, concomitant with decreases in the Firmicutes phyla, and thus an increased Bacteroidetes/Firmicutes ratio, irrespective of diet assignment. Ferrer et al. [24] confirmed these findings comparing the gut microbiota in lean and obese individuals. While many postulated mechanisms of how the gut microbiota contributes to obesity have been suggested [9, 27–30], the focus of this study will be on addressing the relationship between the gut microbiota and SCFAs.

The gut microbiota in and of itself appears to be influenced by many external factors in the host's environment [31], thus when investigating this microbial ecosystem, other influencing external factors must be considered [32, 33]. Previous studies [34, 35], however, are limited by contradictory findings [36], small sample sizes [37–42], imprecise measurements of obesity [43, 44], and lack of detailed dietary and other environmental exposures/mediators [38, 41, 45]. The parent study; Modeling the Epidemiologic Transition Study (NIH R01-DK080763) is a well-established and ongoing prospective cohort study designed to assess the association between body composition, PA, and relative weight, weight gain and cardiometabolic disease risk in five diverse population-based samples of African descent. The five international research sites include Ghana, South Africa, Jamaica, the Seychelles, and the US. The new study, METS-Microbiome (NIH R01-DK111848), is therefore well suited to examine the role a host's local environment has in the associations between the gut microbiota, SCFAs and adiposity. Indeed, each of the 5 METS sites has been well characterized for their own unique environmental and dietary/lifestyle sources of exposure [2, 46–55]. Notably, sites differ according to levels of adiposity (measured using dual x-ray absorptiometry), ranging from 28% in Ghana, up to 39% in the USA, dietary composition (averaged from two 24 h recall), where %diet from carbohydrate ranges 46% in the USA up to 66% in Ghana, as well as differing levels of physical activity (PA, objective activity monitoring).

Using our epidemiologic model for studying the associations between the gut microbiota, SCFAs and the development of obesity, we can explore the interplay of these factors independently and collectively (e.g. dietary habits, daily PA, socio-economic status, public health policy as well as access to health care). In fact, this model has been key to our understanding of obesity and also other chronic diseases in the modern world [4, 54, 56–69]. However, the human gut microbiota, SCFAs and its implications for the obesity epidemic, is only now being considered in detail [37–39, 45]. Interestingly, and to the best of our knowledge, the gut microbiota and SCFAs have not been considered in relationship to the epidemiologic transition model. By exploring these variables through the epidemiologic transition model, we will be able to capture these interactions, and provide novel insight into the obesity epidemic as well as explore innovative therapeutic targets. In fact, we have just published a review justifying the use of this epidemiologic model to unpack the role of the gut microbiota [70].

In summary, the significance of METS-Microbiome is that it may clarify the relationships between gut microbiota, SCFAs and obesity across diverse environments. Also, it may provide novel therapeutic targets, which

might be considered as part of the multi-faceted obesity treatment approach. Specifically, if SCFAs mediate the relationship between gut microbiota and obesity, targeting them through either dietary, probiotics, or pharmaceutical intervention may provide additional therapeutic tools in treatment of obesity.

### METS-microbiome study hypotheses and aims

The METS-Microbiome study was designed to test three hypotheses associated with the relationship between novel risk factors, the gut microbiome and SCFAs, on obesity, adiposity and weight change. We hypothesized that: 1) there exists a shared gut microbiota and SCFAs production are etiological factors in obesity across populations, 2) gut microbiota and SCFAs factors cross-sectionally associated with adiposity will be predictive of longitudinal changes in adiposity, 3) The relationship between gut microbiota and SCFAs production is both shared, yet also reliant on local environmental stimuli.

## Methods

### Design and settings

The parent study; METS, is a well-established and ongoing prospective cohort study designed to assess the association between body composition, PA, and relative weight, weight gain and cardiometabolic disease risk in five diverse population-based samples of African descent (NIH R01-DK080763). A description of the METS protocol for centralized field staff training, data collection, measurement and laboratory procedures has been published [71]. To date, 26 METS-related manuscripts have been published or in press [2, 18, 46–55, 68, 72–74].

In the original METS study, 2,506 ( $N = 2,506$ ) young adults, age 25–45 years, were enrolled at baseline between January 2010 and September 2011 with 500 participants (~50% male) from each of five sites: rural Ghana (Kumasi), peri-urban Republic of South Africa (Cape Town), island nation Seychelles (Mahé), urban Jamaica (Kingston) and suburban Chicago (Maywood, IL) in the United States (USA). These five sites were chosen to represent the spectrum of the ‘epidemiologic transition’ with Ghana and the USA representing the two extremes. Populations sampled represent a range of social and economic development as defined by the United Nations Human Development Index (HDI)(UN [75]). Baseline characteristics of the cohort, are presented by HDI site ranking in Table 1. As a result of the cohort design, average baseline BMI varied widely across sites and obesity ( $BMI \geq 30$ ) prevalence ranged from 1.4% (Ghanaian men) to 63.8% (USA women).

For the new study; METS-Microbiome (R01-DK111848), data and biological samples collected during subsequent years of follow-up examinations will be utilized for year 8–10 follow up (2018–2021). As indicated, a total of 2506 participants were recruited at baseline (2010–2011). At present, across all sites, we have approximately 65% retention of the original cohort. Recruitment and replacement of participants lost to follow-up has begun in the sites.

### Ethics approval

The protocol for METS-Microbiome was approved by the Institutional Review Board of Loyola University Chicago, IL, USA; the Committee on Human Research Publication and Ethics of Kwame Nkrumah University of Science and Technology, Kumasi, Ghana; the Research

**Table 1** Baseline Characteristics of the Original METS Cohort (2506) by Site<sup>a</sup> (mean  $\pm$  SD, %)

	Ghana ( $n = 500$ )	South Africa ( $n = 504$ )	Jamaica ( $n = 500$ )	Seychelles ( $n = 500$ )	USA ( $n = 502$ )
Age (y)	34.3 $\pm$ 6.7	33.4 $\pm$ 5.8	34.4 $\pm$ 6.1	36.1 $\pm$ 5.6	35.3 $\pm$ 6.3
BMI ( $\text{kg}/\text{m}^2$ )	24.1 $\pm$ 4.6	27.5 $\pm$ 8.1	26.5 $\pm$ 6.4	27.1 $\pm$ 5.7	31.9 $\pm$ 8.5
Fat Mass (kg)	18.5 $\pm$ 9.2	29.0 $\pm$ 16.0	24.7 $\pm$ 13.4	25.8 $\pm$ 11.1	37.7 $\pm$ 18.3
Body Fat (%)	28.3 $\pm$ 10.3	36.8 $\pm$ 11.4	31.0 $\pm$ 11.4	33.1 $\pm$ 9.2	39.1 $\pm$ 11.0
Plasma Glucose (mg/dL)	99.7 $\pm$ 12.1	84.0 $\pm$ 23.0	93.1 $\pm$ 9.4	100.7 $\pm$ 29.1	100.1 $\pm$ 34.8
% Carbohydrate (%Energy)	65.8 $\pm$ 10.4	54.9 $\pm$ 11.8	58.5 $\pm$ 8.5	51.2 $\pm$ 9.3	45.9 $\pm$ 9.4
% Protein (%Energy)	11.9 $\pm$ 4.0	16.6 $\pm$ 4.8	14.8 $\pm$ 4.1	18.4 $\pm$ 4.7	15.5 $\pm$ 4.1
% Fat (%Energy)	21.7 $\pm$ 9.2	26.4 $\pm$ 9.3	25.6 $\pm$ 6.6	28.4 $\pm$ 7.6	36.6 $\pm$ 7.0
Fiber (g/day)	25.0 $\pm$ 9.7	8.7 $\pm$ 4.0	16.5 $\pm$ 8.4	13.5 $\pm$ 6.7	14.2 $\pm$ 7.1
Soluble Fiber (g/day)	6.0 $\pm$ 2.8	2.9 $\pm$ 1.4	4.9 $\pm$ 2.7	3.9 $\pm$ 2.1	4.6 $\pm$ 2.4
Insoluble Fiber (g/day)	18.9 $\pm$ 7.5	5.8 $\pm$ 2.9	11.6 $\pm$ 6.1	9.6 $\pm$ 5.1	9.5 $\pm$ 5.38
Obesity (%)	9.8	32.1	27.6	26.4	52.4
Diabetes (%)	1.0	2.4	0.2	3.4	9.6

<sup>a</sup>Sites are presented according to their level of Epidemiologic Transition, using Human Development Index rankings. Diabetes, random glucose > 140 mg/dL or currently taking diabetes medication

Ethics Committee of the University of Cape Town, South Africa; the Board for Ethics and Clinical Research of the University of Lausanne, Switzerland; the Health Research and Ethic Committee of the Ministry of Health of Seychelles, and the Ethics Committee of the University of the West Indies, Kingston, Jamaica. The study strictly adheres to the principles and protocols from the Declaration of Helsinki. The study was registered prospectively with the U.S. National Library of Medicine ClinicalTrials.gov website on December 20, 2017, and began recruiting in January 2018. The study was assigned the following ClinicalTrials.gov identifier: NCT03378765, and is funded by the National Institutes of Health R01 mechanism (R01-DK111848).

### Biological samples and measurements

For METS-Microbiome, 3 years of data will be collected, coinciding with years 8–10 of the original METS study. Project coordinators for each field site were trained and certified in all measurement protocols by coordinating center staff; the measurements included in the METS-Microbiome study are summarized in Table 2. In brief, anthropometrics including weight, height, waist and hip circumferences have been collected using standardized methods and the same equipment [68]. Blood pressure is measured in triplicate at two-time points during each examination using an automatic digital monitor (model HEM-747Ic, Omron Healthcare, Bannockburn, IL USA). Body composition is assessed in all participants at each

examination using bioelectrical impedance analysis and study-specific Eqs. (55). Fasting plasma glucose will be measured; insulin, leptin and adiponectin will be measured in fasting plasma samples using radioimmunoassay kits (Linco Research, Inc., St. Charles, MO). Spot urines will be collected at baseline and assayed for urinary albumin and creatinine levels. Unused whole blood, plasma, serum and urine samples are stored at -80°C for use in future analyses. Fecal samples will be analyzed for both gut microbiota and SCFAs, in all participants from Year 8–10 samples.

### Year 8–10 follow-up examination

All participants, including the original cohort and the new recruits, will undergo the Year 8–10 examination, and as described in the original METS protocol manuscript [49]. Anthropometrics, blood pressure, body composition by BIA, physical activity by accelerometry (Actical; Philips Respironics, Bend OR), and health and medication history by questionnaire will be collected. Extensive information is collected at each examination regarding self-reported health history, focusing on changes to health status since prior visits. Data on drinking, smoking and drug use, prescribed, over-the-counter and illicit, are collected at each examination, along with measures of socioeconomic status, education, employment status and history [76], and physical activity by questionnaire [77]. Site-specific food frequencies will be administered to participants by trained study staff. All

**Table 2** Proposed Study Measures

	Baseline	Follow-Up Examinations		
		Years 1–5	Year 8	Years 9–10
METS				
Physical Activity				
Accelerometer	X	X	X	X
Questionnaire	X		X	X
Body Composition				
BIA	X	X	X	X
Isotope Dilution	X	X		
DXA <sup>a</sup>		X		
Dietary Intake	X		X	
Anthropometrics	X	X	X	X
Serum & Urine Measures <sup>b</sup>	X	X	X	X
Health history, Demographics, etc.	X	X	X	X
METS-Microbiome				
Blood SCFAs	X		X	X
OGTT			X	
Gut Microbiota			X	X

<sup>a</sup>DXA for body composition and bone mineral density measures not available only in Seychelles

<sup>b</sup>Serum measures include fasting glucose, insulin, adiponectin, leptin, lipids, CRP, cystatin C at baseline and glucose, adiponectin and leptin at follow-up examinations

participants will undergo an oral glucose tolerance test (OGTT) to assess glucose tolerance and insulin secretion and sensitivity using Minimal Model analysis. Participants will be asked to provide a fecal sample using a standard collection kit (EasySampler stool collection kit, Alpcos, NH).

#### **Oral glucose tolerance test**

Participants will be instructed to fast overnight and refrain from exercise prior to the test. A standard 75-g OGTT will be performed and blood samples will be drawn at 0, 30, 60, 120 min for subsequent determination of plasma glucose, and serum insulin and C-peptide concentrations. Impaired glucose metabolism will be defined using standard criteria as suggested by the ADA [78] in accordance with the IDF [79] and WHO [80].

#### **Measurement of short chain fatty acids in stool**

SCFAs (acetate, propionate, butyrate, formic acid, and isovaleric acid) will be isolated from 10 mg fecal aliquots, and measured using gas chromatography-mass spectrometry (GC/MS), according to the methods outlined in Moreau et al. [81] and Richardson et al. [82]. Briefly, in an aliquot of 10 mg fecal matter sample (with total protein content analysis by Bradford assay for normalization), add 2-ethylbutyrate internal standard in 0.5 ml water and 0.1 ml concentrated hydrochloric acid, shake 30 min with 1 ml MTBE including methylbutyrate internal standard. Decant MTBE phase, dry over sodium sulfate, derivatize with MTBSTFA at 80 °C for 30 min, inject 1 µl onto a 30 m 0.25 mm, 0.25 µm DB5 durable column in a GC/MS with temperature gradient 50–290 °C, scanning 50–550 Da. Spiked recoveries in fecal matter range from 65 to 110% for formic acid to valerate. All short chain fatty acids had better than 7% within- and between-batch reproducibility and quantification limits < 10 pmol injected onto the column. Fecal samples will be centrally stored at -80°C at Loyola University Chicago after shipment from the field sites.

#### **Measurement of gut microbiota**

**DNA extraction, multiplex 16S allele PCR and sequencing** We will quantify microbiome features from amplicon data using existing pipelines [83] to identify strain-level taxonomic markers for all samples. Microbial DNA will be extracted using the PowerSoil-htp 96-well Soil DNA Isolation Kit (MoBio). The 16S rRNA V4 regions will be PCR-amplified and sequenced using the Illumina HiSeq 2500 platform to generate ~100,000,250 bp paired-end reads per sample [84]. All amplicon sequencing data will be quality filtered and de-multiplexed and then subjected to de novo operational taxonomic unit (OTUs)

picking, and subOTU characterization using DeBlur [85]), via the QIIME platform.

#### **Bioinformatics data analysis**

We will perform a Microbiome Wide Association Study (MWAS; [86]) to determine whether fecal microbial biomarkers are predictive of participant variables. Microbial 16S rRNA diversity will be summarized using Chao1 estimator and Shannon index, and the relative proportions of specific taxa. Significant relationships will be tested using generalized linear modeling. UniFrac distances (between-sample beta-diversity), microbial 16S rRNA diversity (alpha diversity, including evenness) will be correlated against the obesity and SCFA variables using multivariate methods such as principal coordinate analysis (PCoA), Analysis of the Composition of Microbiomes (ANCOM; [87]), and permutational multivariate analysis of variance (PERMANOVA). We will also employ correlative network modeling, including correction of multiple testing, to determine if the network associations (based on relative abundance correlation) differs with population, obesity and SCFA concentrations; such differences in node-level topological features of the network can help with interpreting ecological variability in the stability of each microbiome [88, 89]. We will characterize the modularity of these networks using a random walk approach and link these community structures to sample type via random forest modeling and multinomial logistic regression. Relative abundance of bacterial species will be characterized using sub-operational taxonomic unit level [85]. Based on DESeq2 results [90], logistic models will be fit using patient characteristics and SCFA concentrations as dependent variable and microbiome data as independent variables. Variable selection will be integrated to avoid over-fitting. Classification performance will be evaluated using ROC curve and the 0.632+ bootstrap method [91]. Random Forests will also be applied to determine whether the microbiome is predictive of participant variables [92].

#### **Shotgun sequencing**

We will perform shotgun metagenomic sequencing (20 million reads per sample) to characterize the functional metabolic pathways that may be enriched or depleted in different populations, obesity groupings, or SCFA concentrations. Libraries will be generated using 1 ng of input DNA with the Nextera XT protocol (Illumina), and sequenced on the Illumina HiSeq platform (150 bp × 2, 10 samples per lane, Insert size range = 300 bp to 1200 bp). Raw metagenome reads will be quality trimmed using the neson pipeline [93]. Phylogeny will be assigned to reads using MetaPhlAn [94]. Reads will be assembled using IDBA\_UD [95], and population genomes will be binned using MetaBAT [96]. Single copy



marker gene based copy number variation analysis [97] will be used to estimate completion and intra-species contamination in each genome. Reconstructed genomes will be annotated using RAST [98]. These genomes will be cross-referenced against the 16S rRNA amplicon results and comparative genomes between different strains will be regressed against participant variables (e.g. BMI, SCFAs, etc). Functional genes and metabolic pathways will be identified and statistically analyzed using HUMAN [99], and Hidden Markov Models [100] with DIAMOND [101] and the KEGG database [102]. Relative abundance will be assigned to each KEGG Ortholog detected. Finally, we will apply Predicted Relative Metabolic Turnover [103] to produce a predicted metabolite profile based on pathway reconstruction. The relative proportions of specific taxa, genes, or predicted metabolites associated with outcomes will be tested by regression analysis and generalized linear modeling. UniFrac distances (between-sample diversity) will be correlated with participant variables in principal coordinate analysis (PCoA), permutational multivariate analysis of variance (PERMANOVA).

#### Data management

Loyola University Chicago is the coordinating center for the current study. All data forms, questionnaires and dietary recall instruments are scanned and, along with electronic Actical data files, sent via secure transfer to the data manager at the coordinating center. Scanned forms are coded and double data entered by experienced, trained study staff. A series of logic checks are then performed and, when outliers are encountered, discrepancies are followed up with staff at the appropriate field site.

#### Statistical considerations

##### Overall approach and preliminary analysis

To ensure optimal model selection and protect against model overfitting, cross-validation techniques will be used to develop the models. The entire dataset will be randomly split into a training (60% of data) dataset and a test dataset (40%). Data will be divided using block randomization by site and gender to ensure equal contributions from the five sites. Models from each statistical approach will be fit using the training dataset and associations/predictions will be estimated on the validation set. Optimization of models and variable selection will involve a combination of stepwise selection and AIC criterion via 5-fold cross-validation [104]. The average validation error will also be used to assess model performance among the different statistical approaches and to choose the best approach for modeling this data. Univariate and bivariate summary statistics and distributional plots will be examined for all variables and

appropriate transformations considered. Outliers will be identified using the “letter value” procedure which displays mild and severe outliers at the tails of the distribution [105, 106]. Associations between variables of interest will initially be explored with use of smooth scatter plots for continuous variables and cross-tabulations for discrete variables. Participant characteristics and baseline SCFAs values will be summarized by sex: overall and by site. Additionally, SCFAs concentrations will be examined by participant characteristics representing demographic and lifestyle factors that may be associated with exposure and outcomes of interest. Univariable comparisons will be examined via Pearson’s chi-square tests, Pearson’s correlation coefficients and Student’s t-tests, as appropriate. To account for potential differences in SCFAs levels by site, all modeling will adjust for site (in addition to age and gender) and, when feasible, modeling will be conducted within site to determine site-specific effects. All analyses will be performed using SAS version 9.4 (SAS Institutes, Cary, NC) and computing environment R (R Development Core Team, 2005).

#### Covariates of interest

Based on biological considerations, it will be important to consider the following variables, among others, as covariates of interest in our analyses: age, sex, site, BMI, family or previous history of diabetes or currently taking medications for diabetes or hypertension, blood pressure, nutritional status indicators (e.g., dietary nutrient and specific food intakes), smoking, alcohol use, education, occupation and employment status, marital status, parity (females) and physical activity.

#### Discussion

This study capitalizes upon an existing, extensively well described cohort of adults of African-origin initiated in 2009, with significant variability as a result of the widespread geographic distributions, and therefore variation in the environmental covariate exposures. The METS-Microbiome study will substantially advance the understanding of the role gut microbiota and SCFAs play in the development of obesity and provide novel obesity therapeutic targets targeting SCFAs producing features of the gut microbiota. Specifically Studying unique populations as they span the epidemiologic transition, allows us to investigate several risk factors simultaneously, including environmental co-variables, (e.g. local diet/PA), which have been shown to impact both gut microbiota and SCFAs. In addition, continuing follow-up in a previously established, longitudinal cohort of African-origin adults, and considered high-risk for the development of obesity and metabolic disorders will allow us to capitalize on identifying causal factors.

Notably, we can leverage up to 10 years of detailed phenotype information already collected in the parent study, METS, including yearly weight change, to model the contribution of these on the gut microbiota and SCFAs. METS-Microbiome will utilize observational measures include gold-standard techniques such as DXA body composition and objective PA monitoring, in a large, well characterized population cohort and thus address potential confounding such as diet/PA, as well local environmental stimuli.

The investigative team, with its geographic diversity, existing data, established partnerships, and multidisciplinary expertise, is uniquely positioned to conduct this type of research. With this study, we have enriched the existing METS investigative team of epidemiologists, biostatisticians, nutritionists and exercise physiologists with expertise in endocrinology, microbiology, metabolomics and bioinformatics. This transdisciplinary approach will allow for careful and thorough examination of every phase of the study.

In conclusion, the proposed study will explore the unknown causal mechanisms through which SCFAs mediate the relationship between the gut microbiota and adiposity. As a result of the large and diverse cohort, as well as the comprehensive study design, METS-Microbiome has the potential to uncover several new potential mechanisms involved with development of obesity across populations spanning the epidemiologic transition.

#### Abbreviations

ADA: American diabetes association; ANCOM: Analysis of the composition of microbiomes; BIA: Bioelectric impedance analysis; BMI: Body mass index; GC/MS: Gas chromatography-mass spectrometry; HDL: Human development index; IDF: International diabetes federation; METS: Modeling the epidemiologic transition; OGTT: Oral glucose tolerance test; OTUs: Operational taxonomic unit; PA: Physical activity; PCoA: Principal coordinate analysis; PERMANOVA: Permutational multivariate analysis of variance; SCFAs: Short chain fatty acids; USA: United States of America; WHO: World Health Organization

#### Acknowledgements

The authors would like to acknowledge the site-specific clinic staff members as well as the original 2,500 participants to the METS study.

#### Funding

METS-Microbiome is funded in part by the National Institutes of Health (1R01DK111848).

#### Availability of data and materials

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

#### Authors' contributions

LRD is the overall principal investigator for METS-Microbiome. LRD, JPR, KBA, PB, EVL and TEL are the principal investigators at each of the five study sites. JAG is responsible for the microbiome, and bioinformatics analyses and interpretation. BTL is an endocrinologist with expertise in SCFAs metabolism, diabetes and metabolism. AL is the PI of the previous METS study. LL is the project coordinator. LL and LRD drafted the manuscript. All authors read and approved the final manuscript.

#### Ethics approval and consent to participate

The protocol for METS-Microbiome was approved by the Institutional Review Board of Loyola University Chicago, IL, USA. Approval granted under LU209537. All site protocols have been approved by the local participating institutions; specifically by the Committee on Human Research Publication and Ethics of Kwame Nkrumah University of Science and Technology, Kumasi, Ghana; the Research Ethics Committee of the University of Cape Town, South Africa; the Board for Ethics and Clinical Research of the University of Lausanne, Switzerland; the Health Research and Ethics Committee of the Ministry of Health of Seychelles, and the Ethics Committee of the University of the West Indies, Kingston, Jamaica.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Author details

<sup>1</sup>Public Health Sciences, Stritch School of Medicine, Loyola University Chicago, 2160 S. 1st Avenue, Maywood, IL 60153, USA. <sup>2</sup>Department of Physiology, SMS, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. <sup>3</sup>Institute of Social & Preventive Medicine, Lausanne University Hospital, Lausanne, Switzerland. <sup>4</sup>Ministry of Health, Republic of Seychelles, Lausanne, Switzerland. <sup>5</sup>Research Unit for Exercise Science and Sports Medicine, University of Cape Town, Cape Town, South Africa. <sup>6</sup>Solutions for Developing Countries, University of the West Indies, Mona, Kingston, Jamaica. <sup>7</sup>Microbiome Center, Department of Surgery, University of Chicago, Chicago, IL, USA. <sup>8</sup>Division of Endocrinology, Diabetes, and Metabolism, University of Illinois at Chicago, Chicago, IL, USA. <sup>9</sup>Jesse Brown Veterans Affairs Medical Center, Chicago, IL, USA.

Received: 25 April 2018 Accepted: 24 July 2018

Published online: 06 August 2018

#### References

1. Tataranni PA, Harper IT, Snitker S, Del Parigi A, Vozarova B, Bunt J, Bogardus C, Ravussin E. Body weight gain in free-living pima Indians: effect of energy intake vs expenditure. *Int J Obes Relat Metab Disord*. 2003;27(12):1578–83.
2. Luke A, Bovet P, Plange-Rhule J, Forrester TE, Lambert EV, Schoeller DA, Dugas LR, Durazo-Arvizu RA, Shoham DA, Cao G, et al. A mixed ecological-cohort comparison of physical activity & weight among young adults from five populations of African origin. *BMC Public Health*. 2014;14:397.
3. Dugas LR, Harders R, Merrill S, Ebersole K, Shoham DA, Rush EC, Assah FK, Forrester T, Durazo-Arvizu RA, Luke A. Energy expenditure in adults living in developing compared with industrialized countries: a meta-analysis of doubly labeled water studies. *Am J Clin Nutr*. 2011;93(2):427–41.
4. Luke A, Dugas LR, Ebersole K, Durazo-Arvizu RA, Cao G, Schoeller DA, Adeyemo A, Brieger WR, Cooper RS. Energy expenditure does not predict weight change in either Nigerian or African American women. *Am J Clin Nutr*. 2009;89(1):169–76.
5. Ebersole KE, Dugas LR, Durazo-Arvizu RA, Adeyemo AA, Tayo BO, Omotade OO, Brieger WR, Schoeller DA, Cooper RS, Luke AH. Energy expenditure and adiposity in Nigerian and African-American women. *Obesity (Silver Spring)*. 2008;16(9):2148–54.
6. Hruby A, Hu FB. The epidemiology of obesity: a big picture. *Pharmacoeconomics*. 2015;33(7):673–89.
7. Ley RE. Obesity and the human microbiome. *Curr Opin Gastroenterol*. 2010;26(1):5–11.
8. Kotzampassi K, Giamarellos-Bourboulis EJ, Stavrou G. Obesity as a consequence of gut bacteria and diet interactions. *ISRN Obes*. 2014; 2014:651895.
9. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JL. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*. 2006;444(7122):1027–31.
10. Krebs M, Krssak M, Bernroider E, Anderwald C, Brehm A, Meyerspeer M, Nowotny P, Roth E, Waldhausl W, Roden M. Mechanism of amino acid-



- induced skeletal muscle insulin resistance in humans. *Diabetes*. 2002; 51(3):599–605.
11. Bouchard-Mercier A, Rudkowska I, Lemieux S, Couture P, Vohl MC. The metabolic signature associated with the western dietary pattern: a cross-sectional study. *Nutr J*. 2013;12:158.
  12. Tremaroli V, Backhed F. Functional interactions between the gut microbiota and host metabolism. *Nature*. 2012;489(7415):242–9.
  13. Petriz BA, Castro AP, Almeida JA, Gomes CP, Fernandes GR, Kruger RH, Pereira RW, Franco OL. Exercise induction of gut microbiota modifications in obese, non-obese and hypertensive rats. *BMC Genomics*. 2014;15:511.
  14. Matsumoto M, Inoue R, Tsukahara T, Ushida K, Chiji H, Matsubara N, Hara H. Voluntary running exercise alters microbiota composition and increases n-butyrate concentration in the rat cecum. *Biosci Biotechnol Biochem*. 2008;72(2):572–6.
  15. Queipo-Ortuno MI, Seoane LM, Murri M, Pardo M, Gomez-Zumaquero JM, Cardona F, Casanueva F, Tinahones FJ. Gut microbiota composition in male rat models under different nutritional status and physical activity and its association with serum leptin and ghrelin levels. *PLoS One*. 2013;8(5):e65465.
  16. de Oliveira EP, Burini RC. The impact of physical exercise on the gastrointestinal tract. *Curr Opin Clin Nutr Metabol Care*. 2009;12(5):533–8.
  17. De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poulet JB, Massart S, Collini S, Pieraccini G, Lionetti P. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A*. 2010;107(33):14691–6.
  18. Cardoso I, Bovet P, Viswanathan B, Luke A, Marques-Vidal P. Nutrition transition in a middle-income country: 22-year trends in the Seychelles. *Eur J Clin Nutr*. 2013;67(2):135–40.
  19. Looijer-van Langen MA, Dieleman LA. Prebiotics in chronic intestinal inflammation. *Inflamm Bowel Dis*. 2009;15(3):454–62.
  20. Topping DL, Clifton PM. Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. *Physiol Rev*. 2001; 81(3):1031–64.
  21. Sender R, Fuchs S, Milo R. Revised estimates for the number of human and Bacteria cells in the body. *PLoS Biol*. 2016;14(8):e1002533.
  22. Euzéby JP. List of bacterial names with standing in nomenclature: a folder available on the internet. *Int J Syst Bacteriol*. 1997;47(2):590–2.
  23. Ley RE, Turnbaugh PJ, Klein S, Gordon JL. Microbial ecology: human gut microbes associated with obesity. *Nature*. 2006;444(7122):1022–3.
  24. Ferrer M, Ruiz A, Lanza F, Haange SB, Oberbach A, Till H, Bargiela R, Campoy C, Segura MT, Richter M, et al. Microbiota from the distal guts of lean and obese adolescents exhibit partial functional redundancy besides clear differences in community structure. *Environ Microbiol*. 2013;15(1):211–26.
  25. Patil DP, Dhote DP, Chavan SG, Sultan A, Jain DS, Lanjekar VB, Gangawani J, Shah PS, Todkar JS, Shah S, et al. Molecular analysis of gut microbiota in obesity among Indian individuals. *J Biosci*. 2012;37(4):647–57.
  26. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, et al. Enterotypes of the human gut microbiome. *Nature*. 2011;473(7346):174–80.
  27. Dougherty RM, Galli C, Ferro-Luzzi A, Iacono JM. Lipid and phospholipid fatty acid composition of plasma, red blood cells, and platelets and how they are affected by dietary lipids: a study of normal subjects from Italy, Finland, and the USA. *Am J Clin Nutr*. 1987;45(2):443–55.
  28. Chai W, Conroy SM, Maskarinec G, Franke AA, Pagano IS, Cooney RV. Associations between obesity and serum lipid-soluble micronutrients among premenopausal women. *Nutr Res*. 2010;30(4):227–32.
  29. Saydah S, Bullard KM, Cheng Y, Ali MK, Gregg EW, Geiss L, Imperatore G. Trends in cardiovascular disease risk factors by obesity level in adults in the United States, NHANES 1999–2010. *Obesity*. 2014;22(8): 1888–95.
  30. Hall KD, Sacks G, Chandramohan D, Chow CC, Wang YC, Gortmaker SL, Swinburn BA. Quantification of the effect of energy imbalance on bodyweight. *Lancet*. 2011;378(9793):826–37.
  31. Wang TJ, Ngo D, Psychogios N, Dejam A, Larson MG, Vasan RS, Ghorbani A, O'Sullivan J, Cheng S, Rhee EP, et al. 2-Aminoadipic acid is a biomarker for diabetes risk. *J Clin Invest*. 2013;123(10):4309–17.
  32. Goedecke J, Peer N, Steyn K, Victor H, Levitt NS. Insulin secretion in relation to insulin sensitivity in black South African men and women with increasing age. *Johannesburg: JEMDSA*; 2014. p. 14.
  33. Cox LM, Yamanishi S, Sohn J, Alekseyenko AV, Leung JM, Cho I, Kim SG, Li H, Gao Z, Mahana D, et al. Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. *Cell*. 2014; 158(4):705–21.
  34. Walker AW, Ince J, Duncan SH, Webster LM, Holtrop G, Ze X, Brown D, Stares MD, Scott P, Bergerat A, et al. Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *Isme J*. 2011; 5(2):220–30.
  35. Voreades N, Kozil A, Weir TL. Diet and the development of the human intestinal microbiome. *Front Microbiol*. 2014;5:494.
  36. Bell DS. Changes seen in gut bacteria content and distribution with obesity: causation or association? *Postgrad Med*. 2015:1–6.
  37. Hrydzluszko O, Viant MR. Missing values in mass spectrometry based metabolomics: an undervalued step in the data processing pipeline. *Metabolomics*. 2012;8:S161–74.
  38. Hughes G, Cruickshank-Quinn C, Reisdorph R, Lutz S, Petrache I, Reisdorph N, Bowler R, Kechris K. MSprep—summarization, normalization and diagnostics for processing of mass spectrometry-based metabolomic data. *Bioinformatics*. 2014;30(1):133–4.
  39. Tyakht AV, Kostyukova ES, Popenko AS, Belenikin MS, Pavlenko AV, Larin AK, Karpova IY, Selezneva OV, Semashko TA, Ospanova EA, et al. Human gut microbiota community structures in urban and rural populations in Russia. *Nat Commun*. 2013;4:2469.
  40. Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res Clin Pract*. 2014;103(2):137–49.
  41. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the global burden of disease study 2010. *Lancet*. 2012;380(9859):2095–128.
  42. Segal I, Gagjee PP, Essop AR, Noormohamed AM. Lactase deficiency in the south African black population. *Am J Clin Nutr*. 1983;38(6):901–5.
  43. Vrieze A, Out C, Fuentes S, Jonker L, Reuling I, Kootte RS, van Nood E, Holleman F, Knaapen M, Romijn JA, et al. Impact of oral vancomycin on gut microbiota, bile acid metabolism, and insulin sensitivity. *J Hepatol*. 2014; 60(4):824–31.
  44. Vrieze A, Van Nood E, Holleman F, Salojarvi J, Kootte RS, Bartelsman JF, Dallinga-Thie GM, Ackermans MT, Serlie MJ, Oozeer R, et al. Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology*. 2012;143(4):913–6. e917
  45. Cooksey RC, McClain DA. Increased hexosamine pathway flux and high fat feeding are not additive in inducing insulin resistance: evidence for a shared pathway. *Amino Acids*. 2011;40(3):841–6.
  46. Ettinger AS, Bovet P, Plange-Rhule J, Forrester TE, Lambert EV, Lupoli N, Shine J, Dugas LR, Shoham D, Durazo-Arvizu RA, et al. Distribution of metals exposure and associations with cardiometabolic risk factors in the "modeling the epidemiologic transition study". *Environ Health*. 2014;13:90.
  47. Atiase Y, Farni K, Plange-Rhule J, Luke A, Bovet P, Forrester TG, Lambert V, Levitt NS, Kliethermes S, Cao G, et al. A comparison of indices of glucose metabolism in five black populations: data from modeling the epidemiologic transition study (METS). *BMC Public Health*. 2015;15:895.
  48. Refinetti R, Sani M, Jean-Louis G, Pandi-Perumal SR, Durazo-Arvizu RA, Dugas LR, Kafensztok R, Bovet P, Forrester TE, Lambert EV, et al. Evidence for daily and weekly rhythmicity but not lunar or seasonal rhythmicity of physical activity in a large cohort of individuals from five different countries. *Ann Med*. 2015;47(7):530–7.
  49. Luke A, Bovet P, Forrester T, Lambert EV, Plange-Rhule J, Schoeller DA, Dugas LR, Durazo-Arvizu RA, Shoham D, Cooper RS, et al. Protocol for the modeling the epidemiologic transition study: a longitudinal observational study of energy balance and change in body weight, diabetes and cardiovascular disease risk. *BMC Public Health*. 2011;
  50. Orcholski L, Luke A, Plange-Rhule J, Bovet P, Forrester TE, Lambert EV, Dugas LR, Kettmann E, Durazo-Arvizu RA, Cooper RS, et al. Under-reporting of dietary energy intake in five populations of the African diaspora. *Br J Nutr*. 2015;113(3):464–72.
  51. Sani M, Refinetti R, Jean-Louis G, Pandi-Perumal SR, Durazo-Arvizu RA, Dugas LR, Kafensztok R, Bovet P, Forrester TE, Lambert EV, et al. Daily activity patterns of 2316 men and women from five countries differing in socioeconomic development. *Chronobiol Int*. 2015;32(5):650–6.
  52. Shoham D, Dugas LR, Bovet P, Forrester T, Lambert E, Plange-Rhule J, Schoeller D, Brage S, Ekelund U, Durazo-Arvizu R, et al. Car ownership and

- physical activity across the spectrum of human development: modeling the epidemiologic Transition study (METS). *BMC Public Health*. 2014;15:173.
53. Karalius VP, Harbison JE, Plange-Rhule J, van Breemen RB, Li G, Huang K, Durazo-Arvizu RA, Mora N, Dugas LR, Vail L, et al. Bisphenol A (BPA) found in humans and water in three geographic regions with distinctly different levels of economic development. *Environ Health Insights*. 2014; 8:1–3.
  54. Dugas LR, Bovet P, Forrester TE, Lambert EV, Plange-Rhule J, Durazo-Arvizu RA, Shoham D, Kroff J, Cao G, Cooper RS, et al. Comparisons of intensity-duration patterns of physical activity in the US, Jamaica and 3 African countries. *BMC Public Health*. 2014;14(1):882.
  55. Durazo-Arvizu RA, Camacho P, Bovet P, Forrester T, Lambert EV, Plange-Rhule J, Hoofnagle AN, Aloia J, Tayo B, Dugas LR, et al. 25-Hydroxyvitamin D in African-origin populations at varying latitudes challenges the construct of a physiologic norm. *Am J Clin Nutr*. 2014; 100(3):908–14.
  56. Cooper R, Forrester T, Ogunbiyi O, Muffinda J. Angiotensinogen levels and obesity in four black populations. ICSHIB Investigators. *J Hypertens*. 1998; 16(5):571–5.
  57. Cooper RS, Amoah AG, Mensah GA. High blood pressure: the foundation for epidemic cardiovascular disease in African populations. *Ethn Dis*. 2003;13(2 Suppl 2):S48–52.
  58. Cooper RS, Rotimi CN, Kaufman JS, Owoaje EE, Fraser H, Forrester T, Wilks R, Riste LK, Cruickshank JK. Prevalence of NIDDM among populations of the African diaspora. *Diabetes Care*. 1997;20(3):343–8.
  59. Cooper RS, Wolf-Maier K, Luke A, Adeyemo A, Banegas JR, Forrester T, Giampaoli S, Joffres M, Katarinen M, Primatesta P, et al. An international comparative study of blood pressure in populations of European vs. African descent. *BMC Med*. 2005;3:2.
  60. Kaufman JS, Durazo-Arvizu RA, Rotimi CN, McGee DL, Cooper RS. Obesity and hypertension prevalence in populations of African origin. The investigators of the international collaborative study on hypertension in blacks. *Epidemiology*. 1996;7(4):398–405.
  61. Kaufman JS, Owoaje EE, James SA, Rotimi CN, Cooper RS. Determinants of hypertension in West Africa: contribution of anthropometric and dietary factors to urban-rural and socioeconomic gradients. *Am J Epidemiol*. 1996; 143(12):1203–18.
  62. Kaufman JS, Tracy JA, Durazo-Arvizu RA, Cooper RS. Lifestyle, education, and prevalence of hypertension in populations of African origin. Results from the international collaborative study on hypertension in blacks. *Ann Epidemiol*. 1997;7(1):22–7.
  63. Luke AH, Rotimi CN, Cooper RS, Long AE, Forrester TE, Wilks R, Bennett FI, Ogunbiyi O, Compton JA, Bowsheer RR. Leptin and body composition of Nigerians, Jamaicans, and US blacks. *Am J Clin Nutr*. 1998;67(3):391–6.
  64. Luke A, Rotimi CN, Adeyemo AA, Durazo-Arvizu RA, Prewitt TE, Moragne-Kayser L, Harders R, Cooper RS. Comparability of resting energy expenditure in Nigerians and U.S. blacks. *Obes Res*. 2000;8(5):351–9.
  65. Luke A, Guo X, Adeyemo AA, Wilks R, Forrester T, Lowe W Jr, Comuzzie AG, Martin LJ, Zhu X, Rotimi CN, et al. Heritability of obesity-related traits among Nigerians, Jamaicans and US black people. *Int J Obes Relat Metab Disord*. 2001;25(7):1034–41.
  66. Luke A, Durazo-Arvizu RA, Rotimi CN, Iams H, Schoeller DA, Adeyemo AA, Forrester TE, Wilks R, Cooper RS. Activity energy expenditure and adiposity among black adults in Nigeria and the United States. *Am J Clin Nutr*. 2002; 75(6):1045–50.
  67. Luke A, Durazo-Arvizu R, Rotimi C, Prewitt TE, Forrester T, Wilks R, Ogunbiyi OJ, Schoeller DA, McGee D, Cooper RS. Relation between body mass index and body fat in black population samples from Nigeria, Jamaica, and the United States. *Am J Epidemiol*. 1997;145(7):620–8.
  68. Luke A, Bovet P, Forrester TE, Lambert EV, Plange-Rhule J, Schoeller DA, Dugas LR, Durazo-Arvizu RA, Shoham D, Cooper RS, et al. Protocol for the modeling the epidemiologic transition study: a longitudinal observational study of energy balance and change in body weight, diabetes and cardiovascular disease risk. *BMC Public Health*. 2012;11:927.
  69. Luke A, Adeyemo AA, Tayo B, Durazo-Arvizu RA, Schoeller DA, Leman C, Cooper RS. Energy expenditure, adiposity and weight gain in Yoruba and African-American women. *Obesity Reviews*. 2006;7(Supplement 2):AOP0183.
  70. Dugas LR, Fuller M, Gilbert J, Layden BT. The obese gut microbiome across the epidemiologic transition. *Emerg Themes Epidemiol*. 2016;13:2.
  71. Luke A, Bovet P, Forrester TE, Lambert EV, Plange-Rhule J, Schoeller DA, Dugas LR, Durazo-Arvizu RA, Shoham D, Cooper RS, et al. Protocol for the modeling the epidemiologic transition study: a longitudinal observational study of energy balance and change in body weight, diabetes and cardiovascular disease risk. *BMC Public Health*. 2011;11:927.
  72. Lax S, Smith DP, Hampton-Marcell J, Owens SM, Handley KM, Scott NM, Gibbons SM, Larsen P, Shogan BD, Weiss S, et al. Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science*. 2014;345(6200):1048–52.
  73. Gonseth S, Dugas L, Viswanathan B, Forrester T, Lambert V, Plange-Rhule J, Durazo-Arvizu R, Luke A, Schoeller DA, Bovet P. Association between smoking and total energy expenditure in a multi-country study. *Nutr Metab (Lond)*. 2014;11(1):48.
  74. Luke A, Bovet P, Forrester TE, Lambert EV, Plange-Rhule J, Durazo-Arvizu RA, Kroff J, Richie WN, Schoeller DA. Prediction of fat-free mass using bioelectrical impedance analysis in young adults from five populations of African origin. *Eur J Clin Nutr*. 2013;67(9):956–60.
  75. United Nations Development Programme. Human Development Index (HDI) [<http://hdr.undp.org/en/content/human-development-index-hdi>]. Accessed 10 Jan 2016.
  76. Core Welfare Indicators Questionnaire. The World Bank Group. 2013. <http://go.worldbank.org/66ELZUGJ30>. Accessed 10 Jan 2016.
  77. Armstrong T, Bull F. Development of the World Health Organization global physical activity questionnaire (GPAQ). *J Public Health*. 2006;14(2):66–70.
  78. American Diabetes Association. Standards of medical Care in Diabetes-2014. *Diabetes Care*. 2014;37(Suppl 1):S14–80.
  79. International Diabetes Federation. Clinical Guidelines Task Force: Global Guideline for Type 2 Diabetes. <https://www.idf.org/e-library/guidelines/79-global-guideline-for-type-2-diabetes>. Accessed 10 Jan 2016.
  80. World Health Organization. Screening for type 2 diabetes. In: Report of a World Health Organization and International Diabetes Federation meeting. Geneva; 2003.
  81. Moreau NM, Goupuy SM, Antignac JP, Monteau FJ, Le Bizet BJ, Champ MM, Martin LJ, Dumon HJ. Simultaneous measurement of plasma concentrations and <sup>13</sup>C-enrichment of short-chain fatty acids, lactic acid and ketone bodies by gas chromatography coupled to mass spectrometry. *J Chromatogr B Anal Technol Biomed Life Sci*. 2003;784(2):395–403.
  82. Richardson AJ, Calder AG, Stewart CS, Smith A. Simultaneous determination of volatile and non-volatile acidic fermentation products of anaerobes by capillary gas chromatography. *Lett Appl Microbiol*. 1989;9(1):5–8.
  83. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JL, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7(5):335–6.
  84. Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, Gilbert JA, Jansson JK, Caporaso JG, Fuhrman JA, et al. Improved Bacterial 16S rRNA Gene (V4 and V4–5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems*. 2016;1(1):1.
  85. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, Kightley EP, Thompson LR, Hyde ER, Gonzalez A, et al. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems*. 2017;2(2):1.
  86. Gilbert JA, Quinn RA, Debelius J, Xu ZZ, Morton J, Garg N, Jansson JK, Dorrestein PC, Knight R. Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature*. 2016;535(7610):94–103.
  87. Mandal S, Van Treuren W, White RA, Eggesbo M, Knight R, Peddada SD. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb Ecol Health Dis*. 2015;26:27663.
  88. Cardona C, Weisenborn P, Henry C, Gilbert JA. Network-based metabolic analysis and microbial community modeling. *Curr Opin Microbiol*. 2016;31: 124–31.
  89. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G, et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*. 2017;551(7681):457–63.
  90. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
  91. Efron B, Tibshirani R. Improvements on Cross-Validation: The .632+ Bootstrap Method. *J Am Stat Assoc*. 1997;92(438):548–60.
  92. Starnikov A, Henaff M, Narendra V, Konganti K, Li Z, Yang L, Pei Z, Blaser MJ, Aliferis CF, Alekseyenko AV. A comprehensive evaluation of multicategory classification methods for microbiomic data. *Microbiome*. 2013;1(1):11.
  93. Neson Cookbook [<https://github.com/Victorian-Bioinformatics-Consortium/neson>]. Accessed 1 May 2018.

94. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods*. 2012;9(8):811–4.
95. Peng Y, Leung HC, Yiu SM, Chin FY. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*. 2012;28(11):1420–8.
96. Kang DD, Froula J, Egan R, Wang Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*. 2015;3:e1165.
97. Kerepesi C, Banky D, Grolmusz V. AmphoraNet: the webserver implementation of the AMPHORA2 metagenomic workflow suite. *Gene*. 2014;533(2):538–40.
98. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, et al. The RAST server: rapid annotations using subsystems technology. *BMC Genomics*. 2008;9:75.
99. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B, et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol*. 2012;8(6):e1002358.
100. Prestat E, David MM, Hultman J, Tas N, Lamendella R, Dvornik J, Mackelprang R, Myrold DD, Jumpponen A, Tringe SG, et al. FOAM (functional ontology assignments for metagenomes): a hidden Markov model (HMM) database with environmental focus. *Nucleic Acids Res*. 2014; 42(19):e145.
101. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015;12(1):59–60.
102. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res*. 2008;36(Database issue):D480–4.
103. Larsen PE, Collart FR, Field D, Meyer F, Keegan KP, Henry CS, McGrath J, Quinn J, Gilbert JA. Predicted relative Metabolomic turnover (PRMT): determining metabolic turnover from a coastal marine metagenomic dataset. *Microb Inform Exp*. 2011;1(1):4.
104. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York City: Springer; 2009.
105. Tukey J. Exploratory data analysis. Reading: Addison-Wesley; 1977.
106. Hoaglin D. Letter values: a set of selected order statistics. In: Hoaglin D, Mosteller F, Tukey J, editors. Understanding robust and exploratory data analysis. New York: Wiley; 1983.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

